IJSPT

ORIGINAL RESEARCH

INTER AND INTRA-RATER RELIABILITY OF THE DROP VERTICAL JUMP (DVJ) ASSESSMENT

Emily B. Beyer, PT, DPT, SCS[1*]
Rena F. Hale, PhD[1*]
Aaron R. Hellem, PT, DPT, SCS, OCS, CSCS[1]
Allison M. Mumbleau, PT, DPT, SCS[1]
Nathan D. Schilaty, DC, PhD[1]
Timothy E. Hewett, PhD[2]

ABSTRACT

Background: Non-contact injuries are common in sports as abnormal lower extremity joint mechanics can place athletes at risk for injury. It is important to have reliable, feasible, cost-effective assessment tools to determine lower limb control and injury risk.

Hypothesis/Purpose: The purpose of the study was to assess the intra- and inter-rater reliability of a three-tiered anterior cruciate ligament (ACL) injury risk rating assessment of the drop vertical jump using frontal plane, two-dimensional (2-D) motion capture.

Study Design: Repeated measures.

Methods: Twenty male elite basketball athletes performed the drop vertical jump during a 2-D video assessment at Mayo Clinic Sports Medicine Center in Minneapolis, Minnesota. DVJ scores indicated the following: 1 no visible knee valgus, 2 slight wobble, inward motion of the knees, and 3 knee collision or large frontal plane knee excursion. Score assessment from video of the drop vertical jump was obtained by four independent investigators. The four raters then re-examined the same videos 1 month later, blinded to their original scores.

Results: Intra-rater reliability Fleiss Kappa measure of agreement was substantial amongst all four raters at all scoring time points: initial contact (0.672), first landing (0.728), second landing (0.670), and peak valgus (0.662) ($p < 0.001$). The intra-rater ICC values were good at initial contact (0.809), second landing (0.874), and max valgus (0.885), however were excellent at first landing (0.914) ($p < 0.001$). Inter-rater reliability Fleiss Kappa measurement scores were slight at initial contact (0.173), fair at max valgus (0.343), and moderate at first landing (0.532) and second landing (0.514; $p < 0.001$). Inter-rater ICC values were moderate at initial contact (0.588), excellent at first landing (0.919), and good at second landing (0.883) and max valgus (0.882; $p < 0.001$).

Conclusion: When comparing scores of the drop vertical jump between four independent raters across two sessions, the study demonstrated substantial Kappa and good to excellent ICC intra-rater reliability. Inter-rater reliability demonstrated slight to moderate Kappa measurements of agreement and moderate to excellent ICC's. Thus, for excellent reliability using this assessment, patients should be scored by one individual. For moderate reliability between multiple raters, the first landing of the DVJ should be scored. Findings indicate that the proposed drop vertical jump assessment may be used for reliable identification of abnormal landing mechanics.

Level of Evidence: Level 3

Keywords: Basketball, drop vertical jump, elite athlete, reliability, two-dimensional 2-D, movement screening

*Co-First Authorship
[1]Mayo Clinic Sports Medicine, Minneapolis, MN, USA
[2]Sparta Science, Menlo Park, CA, USA

The protocol for this study was approved by the Mayo Clinic Internal Review Board

The authors have no conflicts of interest to disclose.

CORRESPONDING AUTHOR

Emily B Beyer
beyer.emily@mayo.edu
Mayo Clinic Sports Medicine
600 Hennepin Avenue, Suite #310
Minneapolis, MN 55403

INTRODUCTION

Knee injuries in sports are common, costly, and in some cases preventable.[1] Abnormal lower extremity joint mechanics and altered neuromuscular control can place athletes at increased risk for lower extremity injuries such as an anterior cruciate ligament (ACL) tear.[2-4] Three-dimensional (3-D) motion capture is considered the gold standard for movement screening.[5] However, such testing is time consuming, expensive, requires an expert for data collection and post-processing, and is not available to most clinicians.[2,6,7,8,9] To address these limitations, two-dimensional (2-D) video analysis has been applied to measure frontal and sagittal plane trunk, hip, knee, and ankle kinematics[10,11] and associated lower extremity injury risk factors such as deficits in dynamic stability of the trunk,[12] lower extremity asymmetries[3,13] and valgus at the knee.[5,8,9,14-17] Such testing is used for quick and easy screening with instant feedback for athletes.[5,18,19]

To determine the validity of 2-D analysis as a substitute for lower extremity 3-D analysis, reliability and validity studies have previously been conducted.[11,20,21] Authors have found moderate correlations between 2-D and 3-D measures for the sagittal plane when running[20] and poor correlation for the frontal plane when performing a single leg squat.[11] Previous examiner's intra-rater reliability were excellent (ICCs: 0.951–0.963).[20] However, many clinical studies have used 2-D analysis for assessment purposes rather than 3-D analysis.[5,18,19,24] Such assessments do not require measurements but rather a "yes/no," rating assessment by an experienced reviewer. The assessment screenings are clinically useful since they are not being compared to 3-D analysis thus the measurement of a joint angle or moment is not necessary. The drop vertical jump (DVJ) is a movement tool that is sensitive for detection of increased risk of ACL injury.[22] Researchers have found that increased knee valgus angles and external valgus moments were predictive of future ACL injury; therefore, it has been recommended that an athlete's neuromuscular control be evaluated using dynamic tasks such as the DVJ landing prior to participation in sport.[3,5] The DVJ motion is easy to learn and perform, can be completed in a timely manner for a large number of subjects, and requires little space to administer.[18] However, since previous studies focus on 3-D lower extremity risk assessments and such technology is not readily available at most clinics and current 2-D scoring assessments take more time, more analysis and more assessment, a quicker, easier to administer and analyze 2-D visual scoring assessment is needed for the DVJ motion. The purpose of the study was to assess the intra- and inter-rater reliability of a three-tiered anterior cruciate ligament (ACL) injury risk rating assessment of the drop vertical jump using frontal plane, 2-D motion capture. The hypothesis tested was that intra- and inter-rater reliability would be acceptable for use in the clinic. Results from this study will inform clinicians regarding a feasible, cost-effective, and an expedient assessment tool to access lower limb control and ACL injury risk.

METHODS

Subjects

Subjects consisted of 20 male elite basketball players. All participants gave informed consent to participate in the study which was approved by the Institutional Review Board. The athletes height (1.98 $\pm$ 0.09 m), weight (96 $\pm$ 21 kg), and age (21.7 $\pm$ 1.4 years) were recorded. All were physically active athletes, participating in elite basketball leagues.

Procedure

All athletes performed a 5-minute, self selected intensity warm up on a stationary bike. They wore athletic shoes of their choice provided by the league or their contract and were instructed to perform three DVJ's from a 31 cm high box. Instructions were as follows, "Stand with feet shoulder width apart and toes over the edge of the box. Lean forward until you fall off the box. Do not hop or step off the box but rather fall. Land equally on both feet and immediately perform a maximum vertical jump, reaching with your arms as high as possible as if jumping for a rebound in basketball." Instructions were followed by a demonstration of the jump and emphasis on how quickly they should get off the ground and how high they should reach with both hands. Subjects underwent two to three practice repetitions and then three official jumps were video recorded using 2-D Dartfish camera (Bosch, Stuttgart, Germany) sampling at 60Hz. A camera was placed on a tripod set to the height of the athlete's waist a, two meters

away from the box in the frontal plane. The setup was standardized across all recordings.

DVJ Assessment and Analysis

Four raters independently viewed and scored the DVJ videos. Three of the raters were Doctors of Physical Therapy, Board Certified in Sports Physical Therapy. The final rater was a PhD, researcher in sports medicine and biomechanics. All had five to 10 years of experience. The training procedure was a (Appendix 1) handout to each rater with a five minute explanation of the 3-tiered scoring technique. Raters were instructed to rate knee valgus, using the 3-tiered system, at four time points: 1) initial contact, 2) first landing maximum valgus, 3) second landing maximum valgus, and 4) point of maximum valgus from both jumps. The raters re-examined the same videos one month later, blinded to their original scores.

A fifth, blinded investigator compared and analyzed the results. The scores were analyzed in SPSS (v25, IBM, Armonk, NY) with the Fleiss' Kappa[23-25] and the Intraclass Coefficient Correlation ($ICC_{3,k}$) method to assess the reliability between the scores that the raters assigned for the observed knee valgus. Both analyses were used because Fleiss' Kappa is typically used with categorical variables due to its increased accuracy of the assessment, but ICC is more recognized and interpretable by clinicians. Fleiss' Kappa's report 'level of agreement' and ICC's report consistency. Significance was set at $\alpha < 0.05$. The Fleiss' Kappa methodology calculates the degree of agreement in classification of knee valgus over that which would be expected by chance, where

$$\kappa = \frac{\bar{p} - \bar{p}_e}{1 - \bar{p}_e}.$$

Kappa scores are interpreted as follows: <0 = Poor agreement; 0.01 – 0.2 = Slight agreement; 0.21 – 0.4 = Fair agreement; 0.41 – 0.6 = Moderate agreement; 0.61 – 0.8 = Substantial agreement; 0.81 – 1.0 = Excellent agreement.[23-25] ICC values are interpreted as follows: <0.50 = Poor reliability; 0.5 – 0.75 = Moderate reliability; 0.75 – 0.9 = Good reliability, and >0.90 = Excellent reliability.[24,26]

RESULTS

Intra-rater reliability Fleiss Kappa scores and ICC can be found in Table 1. Intra-rater Kappa levels of agreement reliability were substantial at all four time points: initial contact (0.672), first landing (0.728), second landing (0.670) and peak valgus 0.662). ICC scores (correlations) were excellent at first landing (0.914) and good at initial contact (0.809), second landing (0.874), and peak valgus (0.885). All conditional probabilities and *p*-values are reported in Table 1.

Inter-rater reliability Fleiss Kappa and ICC scores can be found in Table 2. Inter-rater agreement Fleiss Kappa was slight at initial contact (0.173), moderate for first (0.532) and second landing (0.514) and moderate for peak valgus (0.343). ICC's demonstrated moderate correlations at initial contact (0.588), excellent at first landing (0.919), and good at second landing (0.883) and peak valgus (0.882).

DISCUSSION

Biomechanical screening measures are commonly utilized to assess body mechanics that place athletes at risk for injury. Altered neuromuscular control contributes to abnormal lower extremity joint mechanics and can place athletes at risk for ACL injury. 2-D and 3-D motion capture have been utilized to assess the mechanics that place athletes at risk for injury. [5,27,28] Previously designed tools for lower extremity screening have mixed reliability.

Table 1. *Intra-rater Fleiss Kappa levels of agreement and Intraclass Correlation Coefficient (ICC) reliability statistics for the Two-Dimensional (2D) Drop Vertical Jump (DVJ) scores.*

Intra-rater Correlations of the 2D DVJ Scores				
Condition	Fleiss' Kappa (95%CI)	*p*-value	$ICC_{(3,k)}$ (95%CI)	*p*-value
Initial Contact	0.672 (0.453, 0.891)	<0.001	0.809 (0.702, 0.877)	<0.001
First Landing	0.728 (0.570, 0.886)	<0.001	0.914 (0.866, 0.945)	<0.001
Second Landing	0.670 (0.494, 0.845)	<0.001	0.874 (0.803, 0.919)	<0.001
Peak Valgus	0.662 (0.502, 0.821)	<0.001	0.885 (0.821, 0.926)	<0.001

Table 2. *Inter-rater Fleiss Kappa levels of agreement and Intraclass Correlation Coefficient (ICC) reliability statistics for the Two-Dimensional (2D) Drop Vertical Jump (DVJ) scores.*

Inter-rater Correlations of the 2D DVJ Scores				
Condition	Fleiss' Kappa (95%CI)	*p*-value	$ICC_{(3,k)}$ (95%CI)	*p*-value
Initial Contact	0.173 (-0.006, 0.352)	0.058	0.588 (0.186, 0.819)	<0.001
First Landing	0.532 (0.403, 0.661)	<0.001	0.919 (0.840, 0.964)	<0.001
Second Landing	0.514 (0.373, 0.654)	<0.001	0.883 (0.768, 0.949)	<0.001
Peak Valgus	0.343 (0.212, 0.474)	<0.001	0.882 (0.767, 0.948)	<0.001

Some authors have found poor to fair inter- and intra-rater reliability[16,27,28] while others have found excellent inter- and intra-rater reliability, specifically in the tuck jump,[19] drop jump tasks,[5,18] and running assessments.[29] Due to this discrepancy, a study was needed to determine reliability. The purpose of this study was to assess intra- and inter-rater reliability of a three-tiered drop vertical jump ACL injury risk assessment using a 2-D camera. Study results indicated substantial intra-rater agreement/reliability amongst all four raters at all scoring time points and moderate to excellent inter-rater reliability at initial contact and first landing however, kappa agreement was poor to moderate. Agreement is categorical, therefore it is much easier to have disparate results. This assessment was designed for clinicians and trainers to use with patients and athletes in the clinic or on the field. It provides a quick injury risk screen that can be used for injury risk reduction or return to play readiness following injury.

The DVJ ACL risk assessment is a new tool designed for fast athletic screening. Previous studies, though 2-D, are time consuming and require expertise in ACL related injury risk.[5,18,19] The LESS Scoring system requires two cameras and 17 different scored items over the course of three different trials.[18] The Tuck Jump assessment is a quicker test than the LESS as it only requires one camera but still requires scoring of 10 different criteria.[19] The proposed assessment requires one camera, scores four time points, and has intra-rater reliability scores similar to previous 2-D assessments.[5,18] In regards to inter-rater reliability, intra-rater scores were slightly lower however, the proposed assessment is consistent with inter-rater reliability being lower than intra-rater reliability in 2-D analysis studies. [18 30]

This study is unique as it correlated results among four raters, including clinicians and researchers, rather than the more common two clinical raters. It is common in clinical testing and longitudinal research studies to have the same athelete tested by multiple people. By assessing the inter and intra rater reliability with four individuals and clinically assessing with 2D versus 3D motion, the proposed assessment is more practical and less time consuming than previous studies. The results indicate that for excellent reliability, patients should be scored by one individual. For moderate reliability between multiple raters, the first landing of the DVJ should be scored for similar grading or assessment of jump mechanics.

This study is not without limitations. Despite attempts to outline specific time points and standardize rater training, there can be subjectivity in the events of initial contact, first landing, second landing, and peak valgus. However, even with this subjectivity, there was significant reliability between raters ($p < 0.001$). The current study utilized 2-D mechanics. Frontal plane motion observed on 2-D video is not equal to the dynamic knee valgus that can be measured using 3-D techniques.[4] However, Ekegren et al. found that there is an association between 2-D and 3-D analysis which makes 2D analysis of knee valgus worthy of inclusion in preliminary athlete screening.[5] 2-D cameras used for the study collected at 60 hz and many cameras collect at 120hz which may provide more clarity for reviewers. The current study did not include female athletes. Future research is needed to determine repeatability in the female population.

CONCLUSION

The DVJ ACL risk assessment requires minimal equipment and takes less than five minutes to administer on each athlete. It is a fast, easy, and cost effective method to screen athletes in the clinic or on the field. It requires one 2-D camera (collection frequency of at least 60hz) and relies upon the visual scoring of one to four time points. The proposed three-tiered 2D DVJ frontal plane assessment is repeatable and reliable for clinical use. However, for excellent reliability, patients should be scored by one individual over time. For moderate correlations between multiple raters, the first landing of the DVJ should be scored.

REFERENCES

1. Myer GD, Ford KR, Hewett TE. Tuck jump assessment for reducing anterior cruciate ligament injury risk. *Athl Ther Today.* 2008;13(5):39-44.
2. Hewett TE, Myer GD, Ford KR, Slauterbeck JR. Preparticipation physical examination using a box drop vertical jump test in young athletes: the effects of puberty and sex. *Clin J Sports Med.* 2006;16(4):298-304.
3. Hewett TE, Myer GD, Ford KR, et al. Biomechanical measures of neuromuscular control and valgus loading of the knee predict anterior cruciate

ligament injury risk in female athletes: a prospective study. *Am J Sports Med.* 2005;33(4):492-501.

4. McLean SG, Walker K, Ford K, Myer G, Hewett T, van den Bogert AJ. Evaluation of a two dimensional analysis method as a screening and evaluation tool for anterior cruciate ligament injury. *Br J Sports Med.* 2005;39(6):355-362.
5. Ekegren CL, Miller WC, Celebrini RG, Eng JJ, Macintyre DL. Reliability and validity of observational risk screening in evaluating dynamic knee valgus. *J Orthop Sports Phys Ther.* 2009;39(9):665-674.
6. Barber-Westin SD, Noyes FR, Galloway M. Jump-land characteristics and muscle strength development in young athletes: a gender comparison of 1140 athletes 9 to 17 years of age. *Am J Sports Med.* 2006;34(3):375-384.
7. Myer GD, Brent JL, Ford KR, Hewett TE. Real-time assessment and neuromuscular training feedback techniques to prevent ACL injury in female athletes. *Strength Condit Journal.* 2011;33(3):21.
8. McLean S, Walker K, van den Bogert AJ. Effect of gender on lower extremity kinematics during rapid direction changes: an integrated analysis of three sports movements. *J Sci Med Sport.* 2005;8(4):411-422.
9. Sigward SM, Ota S, Powers CM. Predictors of frontal plane knee excursion during a drop land in young female soccer players. *J Orthop Sports Phys Ther.* 2008;38(11):661-667.
10. Dingenen B, Staes FF, Santermans L, et al. Are two-dimensional measured frontal plane angles related to three-dimensional measured kinematic profiles during running? *Phys Ther Sport.* 2018;29:84-92.
11. Schurr SA, Marshall AN, Resch JE, Saliba SA. Two-dimensional video analysis is comparable to 3D motion capture in lower extremity movement assessment. *Int J Sports Phys Ther.* 2017;12(2):163.
12. Zazulak BT, Hewett TE, Reeves NP, Goldberg B, Cholewicki J. Deficits in neuromuscular control of the trunk predict knee injury risk: prospective biomechanical-epidemiologic study. *Am J Sports Med.* 2007;35(7):1123-1130.
13. Pappas E, Carpes FP. Lower extremity kinematic asymmetry in male and female athletes performing jump-landing tasks. *J Sci Med Sport.* 2012;15(1):87-92.
14. Noyes FR, Barber-Westin SD, Fleckenstein C, Walsh C, West J. The drop-jump screening test: difference in lower limb control by gender and effect of neuromuscular training in female athletes. *Am J Sports Med.* 2005;33(2):197-207.
15. Willson JD, Davis IS. Utility of the frontal plane projection angle in females with patellofemoral pain. *J Orthop Sports Phys Ther.* 2008;38(10):606-615.
16. Boden BP, Torg JS, Knowles SB, Hewett TE. Video analysis of anterior cruciate ligament injury: abnormalities in hip and ankle kinematics. *Am J Sports Med.* 2009;37(2):252-259.
17. Hewett TE, Torg JS, Boden BP. Video analysis of trunk and knee motion during non-contact anterior cruciate ligament injury in female athletes: lateral trunk and knee abduction motion are combined components of the injury mechanism. *Br J Sports Med.* 2009;43(6):417-422.
18. Padua DA, Marshall SW, Boling MC, Thigpen CA, Garrett Jr WE, Beutler AI. The Landing Error Scoring System (LESS) is a valid and reliable clinical assessment tool of jump-landing biomechanics: the JUMP-ACL study. *Am J Sports Med.* 2009;37(10):1996-2002.
19. Herrington L, Myer GD, Munro A. Intra and inter-tester reliability of the tuck jump assessment. *Phys Ther Sport.* 2013;14(3):152-155.
20. Maykut JN, Taylor–Haas JA, Paterno MV, DiCesare CA, Ford KR. Concurrent validity and reliability of 2d kinematic analysis of frontal plane motion during running. *Int J Sports Phys Ther.* 2015;10(2):136.
21. Norris BS, Olson SL. Concurrent validity and reliability of two-dimensional video analysis of hip and knee joint motion during mechanical lifting. *Physio Theory Pract.* 2011;27(7):521-530.
22. Ford KR, Myer GD, Brent JL, Hewett TE. Hip and knee extensor moments predict vertical jump height in adolescent girls. *J Strength Condit Res.* 2009;23(4):1327.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977:159-174.
24. Gwet KL. *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters.* Advanced Analytics, LLC; 2014.
25. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physl Ther.* 2005;85(3):257-268.
26. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiro Med.* 2016;15(2):155-163.
27. Chmielewski TL, Hodges MJ, Horodyski M, Bishop MD, Conrad BP, Tillman SM. Investigation of clinician agreement in evaluating movement quality during unilateral lower extremity functional tasks: a comparison of 2 rating methods. *J Orthop Sports Physical Ther.* 2007;37(3):122-129.
28. Krosshaug T, Nakamae A, Boden B, et al. Estimating 3D joint kinematics from video sequences of running and cutting maneuvers—assessing the accuracy of simple visual inspection. *Gait Posture.* 2007;26(3):378-385.

29. Dingenen B, Barton C, Janssen T, Benoit A, Malliaras P. Test-retest reliability of two-dimensional video analysis during running. *Phys Ther Sport.* 2018;33:40-47.

30. Dickson D, Hollman-Gage K, Ojofeitimi S, Bronner S. Comparison of functional ankle motion measures in modern dancers. *J Dance Med Sci.* 2012;16(3):116-125.

APPENDIX 1

DROP VERTICAL JUMP 2-D ASSESSMENT

Procedure:

- To be performed in athletic shoes
- Starting position on top of 31 cm high box
- Instructions:
  "Stand feet shoulder width apart with toes over the edge of the box. Lean forward until you fall off the box. Do not hop off. Once you land, immediately perform a maximum vertical jump reaching with both arms as high as possible, as if jumping for a rebound in basketball."
- Tell the athlete you are looking to see how quickly they can get off the ground and how high they can reach with both hands.
- 2-3 practice repetitions
- 3 official trials
- Video record frontal plane

Rating:

1.) Find the frames that displays initial contact, peak valgus during first landing phase, peak valgus during second landing phase, and peak valgus over the entire motion.
2.) Rate knee valgus you see in each frame based on the key below.
3.) Average the three trials (Trial 1 +Trial 2 +Trial 3)/3
4.) Report the individual trial scores and the average.

2D Video ACL Injury Risk Index	
Score	Description
1	No knee valgus
2	Slight Wobble and/or Inward Motion
3	Knees Collide and/or Large Frontal Excursion

Score 1 **Score 2** **Score 3**